

Export diversification across industries and space: do CIS
countries diversify enough?

Oleksandr Shepotylo

Kyiv School of Economics and Kyiv Economics Institute

Vul. Yakira, 13 Office 326

Kiev, Ukraine 04119

oshepotylo@eerc.kiev.ua

Export diversification across industries and space: do CIS countries diversify enough?¹

Abstract: Despite the importance of industrial and geographical diversification of exports, the literature says little about what a “normal” level of diversification is. This paper takes a step in this direction and develops a methodology to measure a normal level of diversification along industry and space dimensions. The degree of export diversification is computed conditional on country characteristics and bilateral trade costs.

The methodology combines several approaches that recently received attention in the trade literature. First, an industry-level gravity model of exports is estimated using a two-stage estimation procedure that accounts for a sample-selection bias and firm-level heterogeneity. Second, the Hausman-Taylor method is applied for a large panel of countries. Finally, the trade projections are generated out-of-sample.

The methodology is applied to measure the degree of export diversification of the CIS countries. In terms of export potential, the results demonstrate substantial deviations of trade from the levels predicted by the gravity model. All CIS countries except Georgia lag behind the region leaders in terms of the degree of export diversification. In particular, the CIS countries extensively engaged in the export of raw materials have the most concentrated export structure. among all the transition countries in terms of their industrial composition. In terms of geographical diversification, Belarus has the least diversified exports among all transition countries.

Keywords: gravity model, trade potential, trade policy analysis, diversification

JEL categories: F12, F14, F17

¹ I am grateful to Tom Coupe and David Tarr for help and support. I also would like to thank seminar participants at the University of Antwerpen, Kyiv School of Economics, 2008 NARSC conference, 2009 SITE meeting, and 2009 ETSG conference for helpful comments.

Export diversification across industries and space: do CIS countries diversify enough?

1 Introduction

The composition of exports is important for sustainable economic growth. The political economy stresses the negative effect of a so-called “resource curse” (Mehlum et al., 2006; Robinson et al., 2006). That is, the extensive extraction and exportation of raw materials coupled with weak institutions creates a situation where the ruling elite is not interested in establishing property rights protection, which in turn lowers long-run economic growth (Acemoglu et al., 2001). The international trade literature points to various mechanisms that link industrial and export diversification with economic growth. For example, Hausmann et al. (2007) link a higher share of high-quality exports with higher future growth. Koren and Tenreyro (2007) find that low-income countries specialize in fewer and more volatile sectors, which in turn leads to higher aggregate volatility. Imbs and Wacziarg (2003) find a U-shaped association between income per capita and industrial concentration along the development path. In addition, several authors find a positive link between export diversification and economic growth (e.g., Lederman and Maloney, 2003; Herzer and Danzinger, 2006).

Despite the importance of industrial and geographical diversification of exports, the literature says little about what a “normal” level of diversification

is. This paper takes a step in this direction and develops a methodology to measure a normal level of diversification along industry and space dimensions. The degree of export diversification is computed conditional on country characteristics and bilateral trade costs.

The methodology combines several approaches that recently received attention in the trade literature. First, an industry-level gravity model of exports is estimated using a new methodology developed by Helpman et al. (2008). It models the selection into positive trading pairs and heterogeneity of firms and suggests a two-step estimation procedure. At the first stage, the selection into exporters and non-exporters is estimated as a firm level decision, which incurs the fixed costs of exporting. At the second stage, the augmented gravity equation is estimated. Taking into account the selection into trading partners is important because the selection process generates a non-random sample of countries with positive trade flows that could lead to estimation biases if the zeros are ignored. In the estimated sample, the share of zero trade flows is 51 percent, which is a substantial number. Accounting for the heterogeneity of firms is important because, as shown by Melitz (2003), shocks to the variable and fixed costs change the exports at the extensive margins,² which leads to a correlation between the error term and the explanatory variables in the gravity model.

² Extensive margins in the trade literature refer to a number of distribution of exporting firms.

Second, unlike Helpman et al. (2008), the model is developed and estimated at an industry level for a panel of 119 source countries and 147 destination countries for the period 2002–2006 by applying the Hausman and Taylor (1981) procedure. It allows keeping invariant country-specific characteristics while accounting for the correlation between some of the explanatory variables and pair-specific fixed effects. The results demonstrate a substantial variation in estimated coefficients of the gravity equation across industries. In addition, the selection into positive exporters and firm level heterogeneity are important factors that have an impact on the estimation of the gravity model for all industries in the model.

Third, the predictions are calculated out-of-sample (Egger, 2000, 2002) meaning that the Commonwealth of Independent States (CIS) countries' data are not included at the estimation stage, but the CIS country characteristics are used when calculating their trade potential based on the coefficients of the estimated gravity model. This approach was implemented, among others, by McPherson and Trumbull (2008) who estimated the unrealized U.S.-Cuban trade potential. Serlenga and Shin (2007) tested the performance of the Hausman-Taylor method in estimating the gravity equation of bilateral trade flows among 15 European countries from 1960 to 2001 and found that it provides more sensible results than fixed or random effects methods.

Finally, export projections are generated by the gravity model. The export projections allow us to compute so-called normal or potential levels of geographical and industrial diversification that are observed among countries

while controlling for differences in economic development, geographical characteristics and location, trade costs, and industrial capacity.

The method is applied to measure export diversification in nine CIS countries: Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Russia, and the Ukraine.³ Arguably, by mid 00's, the CIS region had become well integrated into world markets and now trades close to its trade potential based on the aggregate level of exports (World Bank, 2005). However, its trade is not well diversified in terms of both trading partners and industry composition. Therefore, a considerable gap between potential and actual trade can be observed when more disaggregated exports are analyzed.

The analysis reveals that along the industry dimension, the exports of Azerbaijan, Kazakhstan, Russia, and the Ukraine—countries that export raw materials—are highly concentrated relative to the level of concentration predicted by the model. These countries trade predominantly in energy resources and metallurgy at the expense of manufactured goods, agriculture, and food products. Georgia, on the other hand, has highly diversified exports. Along the geographical dimension, Georgia, Kazakhstan, and Russia have close to normal levels of diversification, while Belarus has the most concentrated exports among all the transition countries. The analysis of the export levels shows that Belarus, Moldova, and the Ukraine trade below the

³ There are no export data for Tajikistan, Turkmenistan, and Uzbekistan.

potential level with EU countries, while Central Asia and the Caucasus countries trade below the potential level with China and India.

The rest of this paper proceeds as follows. In the second section, we provide a literature review. In the third section, we derive the gravity model at the industry level of aggregation. In the fourth section, we present the data and discuss the empirical strategy of the consistent estimation of the model. In the fifth section, we discuss predicted regional trade flows. Finally, section six concludes and discusses directions for further research.

2 Background

After the collapse of the Soviet bloc, Eastern European and CIS countries experienced an unprecedented growth in trade openness. The trade flows previously administratively directed towards a predominantly intra-bloc trade were reoriented by market forces towards non-bloc countries.⁴ The World Bank (2005) reports that exports from Europe and the Central Asia (ECA) region have tripled and that imports have increased by two and one-half times since the mid-1990s.⁵ The main finding of the World Bank report is that today the openness to trade of countries in the ECA region is largely in line and, in many cases, exceeds the trade openness in other countries with a similar level

⁴ Pelzman (1977) found that the integration of the socialist countries into the Council of Mutual Economic Assistance (CMEA) generated a substantial additional intra-bloc trade at the expense of trade with the rest of the world. He estimated the value of trade creation effect at US\$13.2 billion in 1970.

⁵ The ECA region includes both Eastern European and CIS countries.

of development. However, the adjustments to higher trade openness in CIS countries substantially differ from the Eastern European countries' adjustments:

...in the CIS countries ... the average share of ores, metals, and fuels (oil and natural gas) in total exports increased from 38 percent to 47 percent over the period 1996–2003 ... while there has been substantial change over the course of the transition in the commodity composition and factor intensity of trade by the EU-8 and the SEE economies, relatively little has changed in these regards among the CIS countries, which effectively have been frozen in time. The result is that these countries are not active participants in the evolving international division of labor. The existing composition and factor intensity of exports puts the future growth prospects of the CIS at risk. (World Bank, 2005)

There is an ongoing process of polarization that splits countries within the ECA region into two distinct groups: the first group is Euro-centric (Eastern European countries and Turkey) and the other is Russia-centric (CIS countries). The polarization goes along two dimensions. Geographically, the Eastern European countries diversify their trade away from the ECA region towards “old” Europe, while the CIS countries' trade is increasingly confined within the CIS region. Along the product dimension, there is a common tendency for trade concentration in several product categories and lower trade

diversification for both groups of countries. However, the Eastern European countries are increasingly engaged in intra-industry trade and the trade of manufactured products, and the CIS countries' concentrate exports in primary commodities (World Bank, 2005).

Should the CIS countries worry about the increasing concentration of exports in primary commodities and low geographical diversification of trade? There are two strands of the literature that focus on this question. The first strand looks at links between industrial structure and growth. Imbs and Wacziarg (2003) found a U-shaped association between income per capita and industrial concentration that is robust at different levels of aggregation. At the early stages of development,⁶ when a country relies on production and the export of few products, export diversification helps to protect the country from idiosyncratic terms of trade shocks. Such shocks are less likely to negatively affect trade growth when trade expands due to the larger variety of products rather than when it expands due to an increase in volume of existing products (Bleaney and Greenaway, 2001; Hummels and Klenow, 2005; Felbermayr and Kohler, 2006). There are also important knowledge spillovers that enhance economic growth in a more diversified production environment. Experimenting with the production of new varieties, a developing country learns its comparative advantage in a competitive environment with high

⁶ Imbs and Wacziarg identified the turning point at the level of 13,000 of constant 2000 USD (real GDP per capita in Ireland in 1992), the level that had not been reached by any of the CIS countries by 2006.

uncertainty about the demand for new products (Hausmann and Rodrik, 2003). Higher product variety explains almost a third of intra-country increases in total factor productivity (Feenstra and Kee, 2008). Economic growth depends not only on a larger variety of products, but also on products that are more complex and of better quality. A higher share of high-quality exports is associated with higher future economic growth (Hausmann et al., 2007). Finally, moving away from exporting low value-added primary commodities towards exporting high value-added manufactured products is desirable because of a general trend towards a decline in terms of trade of primary commodities (Athukuruola, 2000).

The second strand of the literature links the composition of industrial output and the development of institutions. Countries that rely heavily on production and the export of raw materials tend to have poor institutions (Mehlum et al., 2006; Robinson et al., 2006), which in turn leads to poor property rights protection and lowers economic growth in the long run (Acemoglu et al., 2001). Therefore, from a political economy point of view, a country that has a less diversified industrial structure skewed towards the extraction of raw materials has fewer incentives to improve its institutions and poorer prospects for sustainable economic growth.

Given the evidence, the geographical and industrial diversification of exports is quite important. Despite its importance, this subject did not get enough attention in the literature that studied the degree of integration of CIS and Eastern European countries into global trade flows. At the beginning of

the transition, most researchers investigated how the reorientation of transition countries towards a market economy would impact the aggregate trade flows (i.e., Wang and Winters, 1991; Hamilton and Winters, 1992; and Baldwin, 1994). Wang and Winters (1991) used a sample of 76 market economies to estimate the gravity equation and project a potential for trade for Eastern Europe and the former Soviet Union. They predicted a substantial increase in trade with industrialized countries, especially West Germany and the United States. More recently, Gros and Gonciarz (1996) found that Eastern European trade responded very quickly to the new regime by reorientation towards EU markets and by 1995 did not differ considerably from that of similar Western European countries. At the same time, Havrylyshyn and Al-Atrash (1998) found that the trade of the former Soviet Union countries still was considerably below its potential by the end of the same period. Recently, Babetskaia-Kukharchik and Maurel (2004) estimated that the CIS countries, Russia in particular, did not trade up to their potential with the EU and that they would gain from joining the WTO and improve their market-oriented institutions. However, their approach, which was based on in-sample projections, makes their findings vulnerable to criticism. That is, perhaps the large deviations of actual trade from that predicted may only show a poor fit of the model rather than deviations from unexploited potential trade.

3 Methodology

A modified version of Helpman et al.'s (2008) model is developed in this section. It explains the mechanism of selection into exporting and non-

exporting firms by modeling the export decisions of heterogeneous firms that differ in their productivity. Exporting is costly due to the fixed costs of exporting, which include setting up a distribution network, adjusting to local preferences, and dealing with country-specific legal requirements. The country-pair-specific fixed costs influence the decision of firms to enter the market; only a subset of firms is productive enough to engage in international trade and to compete in foreign markets. The cut-off point separating exporters from non-exporters varies from one country-pair to another and from one industry to another. Hence, the model generates a pattern of bilateral exports that are industry specific and nonsymmetric for a given country-pair. Thus, the model is able to explain why a majority of firms from a given country may find it profitable to export to one destination and not to export at all to another destination due to the country-pair specificity of the fixed costs.

The methodology is different from that of Helpman et al. (2008) in several important ways. First, it is a model at the industry level that allows for industry-level heterogeneity in trade costs.⁷ Second, the source of uncertainty in the model comes from unobservable factors in the multilateral resistance

⁷ Hummels (1999) studied trade costs for 3,000 goods for New Zealand and Latin American imports and over 15,000 goods for U.S. imports and found that trade costs vary significantly across industries. In particular, freight costs for manufacturing are lower than for commodities and agricultural products. For example, importing fruits and vegetables costs approximately 15 percent of the value of shipment, while importing road vehicles costs 2.1 percent.

term in addition to unobservable factors in the trade costs. The unobservable country- and pair-specific factors can be correlated with some of the explanatory variables, and the resulting endogeneity is controlled for by applying the Hausman-Taylor (1981) method that exploits time and cross-country variations in the data. The use of panel data instead of cross-sectional analysis allows us to remove some biases stemming from unobserved industry and country-pair heterogeneity and to estimate the parameters of the model with greater precision.

3.1 Parameterization and Estimation Strategy

The theoretical model, which is presented in the Appendix, allows us to derive the two estimated equations. We estimate the gravity equation as:

$$X_k^{ij} = s_k^i Y^i Y^j V_k^{ij} \frac{\left(\frac{T_k^{ij}}{P_k^j}\right)^{1-\sigma}}{\sum_{j=1}^C \left(\frac{T_k^{ij}}{P_k^j}\right)^{1-\sigma} Y^j V_k^{ij}}, \quad (1)$$

given the condition

$$\pi_k^{ij}(a_k^{ij}) = 0 \Leftrightarrow \frac{\theta_k}{\sigma} \left[\frac{\sigma T_k^{ij} w^i a_k^{ij}}{(\sigma-1) P_k^j} \right]^{1-\sigma} Y^j = F_k^{ij}. \quad (2)$$

In the following discussion, the industry index k is dropped for the reader's convenience.

3.1.1 Selection of firms

A distribution of productivities or the productivity of the marginal exporter a_t^{ij} is not observable. However, we do have information on trading and non-trading country-pairs. We define a latent variable as:

$$\Psi_t^{ij} = \frac{\frac{\theta}{\sigma} \left[\frac{\sigma T^{ij} w_t^i a_{\min}^i}{(\sigma-1) P_t^j} \right]^{1-\sigma} Y_t^j}{F_t^{ij}}. \quad (3)$$

Ψ_t^{ij} is defined as the ratio of the variable profits from exporting to the fixed cost of exporting for the most productive firm. A positive export is observed if $\Psi_t^{ij} \geq 1$. Furthermore, conditional on a positive export, V_t^{ij} is an increasing function of Ψ_t^{ij} for an arbitrary $G(a)$ (see Helpman et al., 2008).

Trade costs associated with the shipping of a unit of good from country i to country j are modeled by assuming the commonly used functional form:

$$(T_t^{ij})^{\sigma-1} = (dist^{ij})^\rho \exp(Z^{ij}\gamma - u_t^{ij}),$$

where $dist^{ij}$ is the distance between countries i and j and Z is a set of additional variables that determine trade costs, such as the contiguity dummy,

whether country i or j is landlocked, the interior distances of countries i and j , and whether the countries are located on the same continent. γ is the vector of coefficients associated with Z . Finally, u_t^{ij} is an unobservable shock to the trade costs distributed $N(0, \sigma_u^2)$ and independent across time periods. Suppose further that fixed costs have the following functional form:

$F_t^{ij} = \exp(\phi^i + \phi^j + \kappa\phi_k^{ij} - \theta_t^{ij})$, where ϕ^i represents fixed costs specific to the exporting country, ϕ^j represents fixed costs specific to the importing country, ϕ^{ij} represents country-pair-specific fixed costs, and θ_t^{ij} represents country-pair-specific random components distributed as $N(0, \sigma_\theta^2)$.

Taking logs of both sides of equation (3) and substituting expressions for the variable and fixed trade costs yields:

$$\psi_{kt}^{ij} = \zeta_0 + (1 - \sigma) \ln w_t^i - \rho \ln dist^{ij} - Z^{ij} \gamma + d_t^j - \phi^i - \phi^j - \kappa\phi^{ij} + \omega_t^{ij}, \quad (4)$$

where $d_t^j = \ln Y_t^j + \phi^j - (1 - \sigma) \ln P_t^j$ is a country j time-specific dummy and

$$\omega_t^{ij} = u_t^{ij} + \theta_t^{ij} \sim N(0, \sigma_u^2 + \sigma_\theta^2).$$

Without loss of generality, we can assume that $\sigma_u^2 + \sigma_\theta^2 = 1$ to normalize the selection equation that brings the following probit model:⁸

⁸ Alternatively, both sides of (9) are divided by σ_η . Both procedures lead to the same outcome in terms of predicting the probability of positive trade.

$$\rho_t^{ij} = \text{Prob}(X_t^{ij} > 0 \mid \Psi_t^{ij}) = \Phi(\zeta_0 + (1 - \sigma) \ln w_t^i - \rho \ln \text{dist}^{ij} - Z^{ij} \gamma + d_t^j - \phi^i - \phi^j - \kappa \phi^{ij}) \quad (5)$$

We denote a predicted probability of positive exports from country i to country j at time t as $\hat{\rho}_t^{ij}$.

Equation (5) is estimated using the standard probit model that includes the destination country time-specific fixed effects. For better identification, several variables that affect fixed costs but that have no effect on the volume of trade (and thus are included in the selection equation but not the gravity equation) are needed. Based on the results from Helpman et al. (2008) and Martin and Pham (2008), we control for pair-specific fixed costs by including a common language dummy as one of the variables that affects the decision of a firm to trade but has no significant impact on the volume of trade. To control for the exporter country-specific fixed costs, we include regulatory quality indices for the reporting country as factors that are proportional to the fixed costs of trade and therefore belong to the selection equation.⁹

3.1.2 Gravity equation

Taking the logs of both sides of equation (1) and substituting for T_t^{ij} yields:

⁹ The regulatory quality index from governance matters (Kaufmann, Kraay, and Mastruzzi, 2007) measures the ability of the government to formulate and implement sound economic policies that promote private sector development.

$$\ln X_t^{ij} = \ln s_t^i + \ln Y_t^i + \ln Y_t^j - \rho \ln dist^{ij} - Z^{ij} \gamma_k + \ln V_t^{ij} - (1 - \sigma) \ln P_t^j - \ln MRT_t^j + u_t^{ij}, \quad (4)$$

where $\ln MRT_t^j = \ln \left(\sum_{j=1}^C \left(\frac{T_t^{ij}}{P_t^j} \right)^{1-\sigma} Y^j V_t^{ij} \right)$ is the multilateral resistance term, an

integral measure of trade barriers of a country vis-à-vis all its trading partners (Anderson and van Wincoop, 2003), which accounts for the endogenous and simultaneous determination of trade flows across all countries. The

multilateral resistance term $\ln MRT_k^i$ is not observable and according to theory is simultaneously determined for all countries. A traditional approach to deal with the multilateral resistance term is by introducing country fixed effects or pair fixed effects (see Baldwin and Taglioni, 2006, for a discussion on the usage of fixed effects in the gravity equation). However, it limits the ability of the model to generate out-of-sample predictions because of the inability to estimate country fixed effects for exporting countries not included in the estimation sample. We assume that the MRT can be approximated—if parameter σ is close to two and that distance contributes the most to the trade

costs—by the following expression: $\ln MRT_t^j = \ln \left(\sum_{j=1}^C \left(\frac{P_t^j Y^j}{dist^{ij}} \right) \right)$. Recognizing

the fact that approximation is far from exact and that the multilateral resistance term is endogenous, we apply the Hausman-Taylor (1981) method that takes into account the endogeneity of the MRT.

Following the semiparametric version of the Helpman et al. (2008) method, we use the information acquired at the first stage of the estimation by identifying $E(\ln V_t^{ij} + u_t^{ij} \mid X_t^{ij} > 0) = b_0 \hat{\eta}_t^{ij} + \sum_{m=1}^3 b_m (\hat{\psi}_t^{ij})^m$, where $\hat{\eta}_t^{ij}$ is the traditional inverse Mills ratio that accounts for the sample selection bias and the polynomial of degree 3 in $\hat{\psi}_t^{ij} = \hat{\eta}_t^{ij} + \Phi^{-1}(\hat{\rho}_t^{ij})$ corrects for the firm level heterogeneity. As shown by Helpman et al. (2008), the polynomial of degree 3 is a sufficiently flexible and accurate approximation of the underlying unknown function of the distribution of productivities $G(a)$.

4 Application: Trade Potential of CIS Countries

4.1 Dependent Variable

The investigated sample includes 128 source countries and 147 destination countries for the period 2002–2006. The export data were acquired from the United Nations Commodity Trade Statistics database (COMTRADE) at the level of Global Trade Analysis Project (GTAP) sectors and is further aggregated into 10 industries according to the mapping presented in Table 1.¹⁰ Potentially, 18,688 positive bilateral exports per industry per year can be observed. However, as shown in Table 2, zeros account for more than one-half

¹⁰ We acquired data for 42 GTAP sectors, excluding service sectors. Further aggregation to 10 industries is done for ease of presentation but is not necessary from theoretical and computational standpoints.

of all the observations. The sample has 455,906 positive trade values out of a possible 934,700 bilateral pairs, which accounts for only 49 percent of the sample. The share of positive exports varies considerably across industries—from 29 percent of positive flows in energy resources to 59 percent of positive flows in the other manufacturing industry, as presented in the last column of Table 2.

Looking at intensive margins of trade, the average value of bilateral exports varies significantly across industries as well. The average export value is equal to US\$24.6 million in the agriculture and forestry industry, US\$73.9 million in metallurgy, and US\$147.9 million in electronic equipment, as reported in the second column of Table 2. Overall, the average value of exports at the industry level is equal to US\$91.2 million.

4.2 Independent Variables

Data on the industrial composition of GDP in exporting country i at time t is not directly available. We use data on total exports of sector k from country i excluding bilateral exports to country j to construct the closest available proxy

$$s_{kt}^i = \frac{\text{export}_{kt}^i - \text{export}_{kt}^{ij}}{\text{export}_t^i}, \text{ which takes into account the time variation in the}$$

composition of industrial output. A bilateral export between countries i and j is excluded to deal with the endogeneity of the sector export share. The suggested proxy would represent the output structure of country i reasonably

well if $\frac{Y_{kt}^i}{Y_t^i} \sim \frac{\text{export}_{kt}^i}{\text{export}_t^i}$. The suggested proxy overestimates an output share of

an export-oriented industry, especially for large countries. However, it works reasonably well for small, open economies that sell most of their output abroad. Also, the measurement error is less of a problem due to the fact that it is industry specific, and we run the model for each industry separately.

GDP in current U.S. dollars and population data were acquired from the 2007 World Development Indicators (WDI). Geographical characteristics and distances between countries were collected from the Centre D'Etudes Prospectives et D'Informations Internationales (CEPII). An interior distance was measured as the average distance within a country, and landlocked dummies were chosen to control for trade costs within the source and destination countries. A contiguity dummy (whether one of the countries in the country-pair was ever a colony of the other country and whether countries are located on the same continent) was used to control for pair-specific trade costs that are not directly related to distance.

4.3 Selection Variables

We chose two variables that enter the selection equation, but not the gravity equation, based on the results of Helpman et al. (2008) and Martin and Pham (2008).¹¹ The common language dummy is the variable that controls for the

¹¹ Martin and Pham (2008) employed a Monte Carlo simulation and demonstrated that ignoring the sample selection problem in the gravity equation (9) leads to substantial biases. They compared various estimation methods, such as truncated OLS, Maximum Likelihood (ML), the Poisson Pseudo Maximum Likelihood (PPML) estimator recommended by Silva

pair-specific fixed costs. It captures fixed costs related to adapting to cultural and linguistic barriers between two countries (differences in religious beliefs, translation, advertising, etc.).

To control for country-specific fixed costs related to institutional quality in exporting countries, we used governance indicators of regulatory quality acquired from Kaufmann et al. (2007). They capture the effectiveness of bureaucracy, amount of red tape, and quality of policies and regulations that encourage free trade and promote private-sector development.

5 Results

This section has the following goals. First, it reports results from the two-stage estimation of the gravity equation. Second, it predicts export flows that are interpreted as potential exports under the assumption that CIS countries are not too different from a typical country included in the sample and compares them with the actual trade flows. Third, it computes indices of industrial and geographical diversification of exports for all CIS countries. Finally, it compares the degree of export diversification in CIS countries with the degree of export diversification in Eastern European countries.

and Tenreyro (2006), Nonlinear Least Squares (NLS), and Heckman's Maximum Likelihood (HML), and found that HML produces reliable estimates with small biases and small standard errors.

5.1 Two-stage Estimation of the Gravity Equation

5.1.1 Selection equation

Table 4 presents the results of the probit regression for each industry estimated for the entire sample of 128 exporting countries and 147 destinations countries from 2002 to 2006. Importantly, variables that appear only in the selection equation are significant and have the coefficients of the expected sign. Sharing a common language increases the probability of positive exports by 17 percent in the agriculture and forestry and timber industries, by 16 percent in the wood and paper industry and the electronic equipment industry, and by 15 percent in the food industry, while a common language is less important for the probability of positive trade in the energy resources industry and the chemical products. Better regulatory quality in an exporting country also improves the chances of positive exports. A one standard deviation increase in the regulatory quality (RQ), which roughly corresponds to institutional differences between Russia (RQ=-0.51 in 2004) and Poland (RQ=0.64 in 2004), increases the probability of positive trade in the electronic equipment industry by 29 percent, in the food industry by 27 percent, and in the motor vehicles and parts industry by 21 percent.

All variables that enter the gravity equation in the second stage are also important determinants of the probability of positive exports. Country i is more likely to export to country j in industry k when i is more economically developed, when it is large in size, and when it has access to the sea. Among country-pair characteristics, positive trade is more likely between countries

that are close to each other, share a common border, are located on the same continent, and share a colonial past. Countries that have common borders are more likely to trade; for example, trade is 23 percent higher in the food industry, 20 percent in the electronic equipment industry, and 18 percent in the energy resources industry. Past colonial relationships have a positive impact on the probability of trade; it ranges from 9 to 26 percent across all industries except for other manufacturing where it has an insignificant effect.

In addition to the control variables reported in the first column of Table 4, regressions include the destination country time-specific fixed effects. Since we use a nominal GDP and nominal export values, time dummies are included to account for common time shocks and to make observations from different time periods comparable (Baldwin and Taglioni, 2006), while destination-country fixed effects correct for variations in price levels. The standard deviations reported in parentheses are cluster robust. Pseudo R-squared, reported at the bottom of Table 4, ranges from 0.31 to 0.41 and shows that selected variables explain the probability of export reasonably well.

5.1.2 Gravity model of bilateral exports corrected for selection and firm-level heterogeneity

Table 5 reports the results of the evaluation of the gravity equation for each industry estimated on the sample of 119 source countries and 147 destination countries for 2002 to 2006. The nine CIS countries are not included at this stage of the estimation. We allow for the endogeneity of the multilateral resistance term or so-called measure of remoteness, and we control for

correlations between remoteness and unobserved random effects u^{ij} by employing the Hausman-Taylor (1981) method that fits panel-data random-effect models in which some of the explanatory variables are correlated with individual-level unobserved heterogeneity. Year and destination fixed effects are included but not reported in Table 5.

The coefficients of the log of export share, the log of GDP i , and the log of GDP j are positive and significant for all sectors, as expected from the theoretical model. At the same time, there is substantial variation in coefficients across industries. It justifies the choice of running a separate regression for each industry rather than a pooled regression with industry fixed effects.¹² The log of distance between countries enters negatively and has substantial cross-industry variability ranging from -0.81 for the electronic equipment industry to -1.64 for the chemical products industry. Exports for most industries are increased when there is a common border, same continent location, and common colonial past. The coefficients for interior distances have a positive sign for agriculture and resource-extracting industries and negative signs for manufacturing industries, which reflect two opposite forces in play—higher transportation costs within a country associated with a larger country size would tend to reduce trade, while a larger country size would increase supply and production of goods that intensively use land and natural resources. Landlocked countries tend to trade less due to higher transportation

¹² A formal econometric test on poolability of the data is strongly rejected

costs (Hummels, 1999; Limao and Venables, 2001). Remoteness of the exporting country is mostly positive for manufacturing industries reflecting the increase in trade between two countries when all other potential trading partners are located farther. At the same time, the remoteness term is negative and significant for the agriculture, food, energy resources, timber wood and paper industries.

The inverse Mills ratio $\hat{\eta}_{kt}^{ij}$ and the polynomial in $\hat{\psi}_{kt}^{ij}$ (variables that approximate $\ln V_{kt}^{ij}$ and control for selection in exporters and firm-level heterogeneity) are jointly significant, as reported at the bottom of Table 5, which shows $\chi^2(4)$ statistics and the corresponding p-value for the test that all coefficients of the approximating polynomial $b_{k0}\hat{\eta}_{kt}^{ij} + \sum_{m=1}^3 b_{km}(\hat{\psi}_{kt}^{ij})^m$ are jointly equal to zero. The sound rejection of the test for all industries indicates the importance of the first-stage selection process in exporters and firm-level heterogeneity on the intensive margins of trade.

5.2 Industrial and Geographical Composition of Exports of the CIS Countries

The developed model allows us to project the results of the estimation procedure on the sample of CIS countries along industry and space dimensions. We generate predicted exports of CIS countries by applying the characteristics of CIS countries to the coefficients of the gravity equation (10), which was estimated in the previous section. We refer to the predicted export

flows as potential export flows \hat{X}_{kt}^{ij} and compare them with the actual exports X_{kt}^{ij} .

We cannot directly compare the levels of actual and predicted exports due to the log-linearized nature of the gravity equation and the fact that, as an illustration of Jensen's inequality, the log of expected value is not equal to the expected value of the log (Silva and Tenreyro, 2006). However, we can compare the shares. To illustrate this point, consider the following example.

Suppose that $\ln X_{kt}^{ij} \sim N(a, \sigma_{ij}^2)$. Calculating the expected value, it can be shown that $\hat{X}_{kt}^{ij} = EX_{kt}^{ij} = \exp(a_{kt}^{ij} + \sigma_{ij}^2 / 2)$. Notice that we cannot compute the level of \hat{X}_{kt}^{ij} without knowing or consistently estimating σ_{ij}^2 , but we can

compute the share $\hat{x}_{kt}^{ij} = \frac{\hat{X}_{kt}^{ij}}{\sum_{i \in I} \hat{X}_{kt}^{ij}} = \frac{\exp(a_{kt}^{ij})}{\sum_{i \in I} \exp(a_{kt}^{ij})}$. Therefore, the following

analysis is carried out by comparing the actual and predicted exports in shares rather than in levels.

Table 6 reports the geographical distribution of actual and predicted exports for CIS countries with selected trading partners. For each country-pair, the top number is $x^{ij} = \frac{X^{ij}}{X^i}$, where X^{ij} is the total actual export of country i to region j and X^i is the total actual export of country i . The bottom number is

$\hat{x}^{ij} = \frac{\hat{X}^{ij}}{\hat{X}^i}$, where actual exports are replaced by potential exports. By

construction, each row of Table 6 adds up to one.

According to the results, CIS countries from the European region (Belarus, Moldova, and the Ukraine) export below the potential level to the EU and above the potential level to other CIS countries. The CIS countries of Central Asia, the Caucasus region, and Russia, on the other hand, tend to overperform in trade with EU countries and considerably underperform in trade with China and to some extent with India. Azerbaijan, Kazakhstan, and Russia export considerably more to EU countries at the intensive margins than predicted by the gravity equation. This pattern is easily explained if we look at the industrial structure of exports reported in Table 7. The above-mentioned countries have an industrial structure of exports that is extremely skewed towards exporting energy resources to EU countries. During the investigated period, the exportation of energy resources accounted for 66 percent of the total export of Kazakhstan; in Azerbaijan and Russia, those numbers are 64 and 44 percent, respectively. There is over-performance by some countries (the Ukraine in particular) in the export of metals. Conversely, there is large, consistent underperformance of potential trade in agriculture and food products in all the CIS countries. For example, according to the gravity model, the Ukraine should export 27 percent of its total exports in agriculture and food, twice as much as it actually exports. Finally, Azerbaijan, Kazakhstan, and Russia—countries that intensively export energy resources—considerably underperform in the export of manufacturing products. For example, the share of exports of motor vehicles and parts, electronic equipment, and other

manufacturing in the total export of Kazakhstan is 2 percent, while the gravity model predicts 11 percent.

Based on the results, we make the following preliminary conclusions. First, there are substantial deviations in trade among Azerbaijan, Kazakhstan, Russia, and the Ukraine from the levels predicted by the gravity model. In particular, their export structure is skewed towards exporting energy resources and metals and away from exporting agriculture, food, and manufactured goods. Second, given the rapid development of China and India, Central Asia and the Caucasus region have surprisingly weak trade relations with those countries. Finally, the CIS countries located close to the EU (Belarus, Moldova, and the Ukraine) trade with the EU below the potential level.

5.3 Industrial and Geographical Diversification of Exports of the CIS Countries

To quantify the degree of diversification, we apply the Herfindahl-Hirschman Index (HHI), which is a traditional measure of industry concentration, to measure the degree of export diversification along industry and space dimensions. For example, to measure the degree of export diversification along trading partners of CIS country i in sector k , we compute the normalized HHI for actual exports:

$$H_k^i = \frac{\sum_{j=1}^{147} (x_k^{ij})^2 - 1/(147*5)}{1 - 1/(147*5)}, x_k^{ij} = \frac{\sum_{t=2002}^{2006} X_{kt}^{ij}}{\sum_{i=1}^9 \sum_{t=2002}^{2006} X_{kt}^{ij}}.$$

\hat{H}_k^i is computed similarly using the predicted exports. Finally, we compute the ratio \hat{H}_k^i / H_k^i and report it in Table 8. $\hat{H}_k^i / H_k^i < 1$ indicates geographical overconcentration of the actual exports relative to predicted exports, while $\hat{H}_k^i / H_k^i > 1$ indicates under concentration.

Overall, the CIS countries' exports are not sufficiently geographically diversified because the average ratio of the two HHI indices for the aggregated exports is equal to 0.72. Across different industries, the light industry, the timber industry, the wood and paper industry, the electronic equipment industry, the motor vehicles and parts industry, and other manufactured goods industry are the least geographically diversified. The highest geographical diversification is observed in industries that comprise the major share of total exports like the energy resources industry in Russia or the agriculture and metallurgy industries in the Ukraine. Of all the CIS countries, Kazakhstan has the most geographically diversified exports. Its index of geographical concentration for the predicted exports exceeds the index for the actual exports by 1.25 times. Kazakhstan's exports are well diversified in the energy resources industry (exceeds predicted level of diversification by 3.25 times), the food industry (by 4.09 times), the chemical industry (by 3.61 times), and

the metallurgy industry (by 2.20 times). Russia's and Georgia's exports are also quite diversified with ratios of 1.01 and 0.99, respectively. Azerbaijan, Armenia, and especially Belarus, on the other hand, have exports that are not geographically diversified. Belarusian exports are overconcentrated primarily because 40 percent of the exports went to Russia over the investigated period. For example, its predicted level of diversification in the food industry exceeds its actual one by 10 times!¹³

To measure the degree of industry diversification of exports of CIS country i at time t , we compute:

$$H_t^i = \frac{\sum_{k=1}^{10} (x_{kt}^i)^2 - 1/(10)}{1 - 1/(10)}, x_{kt}^i = \frac{\sum_{j=1}^{147} X_{kt}^{ij}}{\sum_{k=1}^{10} \sum_{j=1}^{147} X_{kt}^{ij}}. \text{ By analogy, } \hat{H}_t^i \text{ is computed, and the}$$

ratio \hat{H}_t^i / H_t^i is reported in Table 9. Of all the CIS countries, Georgia is the only country that has an industrially diversified structure of exports that has become even more diversified over time. Belarus follows in second place but demonstrates a negative tendency towards higher concentration, while Moldova reached a normal level of diversification by 2006. All other countries have exports that are overconcentrated. Azerbaijan has the most concentrated exports, which exceed the level of concentration predicted by the gravity model by almost five times. Other countries with industrially concentrated

¹³ A recent trade dispute between Belarus and Russia over the quality of dairy products illustrates the vulnerability of Belarusian exports in this respect.

exports include Kazakhstan, Russia, and the Ukraine. It comes as no surprise because those are the countries that are heavily engaged in the export of raw materials. Finally, there is no common tendency in the dynamics of the level of diversification for the CIS countries as a group from 2002 to 2006.

5.4 Export Diversification: Comparison with Other Eastern European Countries

Do the CIS countries systematically differ from other transition countries in terms of industrial and geographical diversification of their exports? Figure 1 plots a degree of relative export diversification along the two dimensions. Countries that are located in the first quadrant of the graph (Czech Republic, Poland, and Slovenia) have exports that are diversified well above normal levels in both dimensions. Georgia has the highest level of industrial diversification among all transition countries with Croatia and the Czech Republic coming in second and third, respectively. Bulgaria and Kazakhstan are geographically diversified above normal levels. Four CIS countries (Azerbaijan, Kazakhstan, Russia, and the Ukraine) are the worst performers in export diversification along the industrial dimension, while Belarus is the worst performer in geographical diversification. Other transition countries' export diversification tends to be slightly less than expected; they are located in the third quadrant relatively close to the origin.

6 Conclusions

This paper estimates a disaggregated gravity equation using a two-stage procedure and computes the export potentials of CIS countries at the industry level of aggregation. It further calculates the HHI, a traditional measure of industry concentration, to measure the degree of export diversification along industry and space dimensions. Four CIS countries (Azerbaijan, Kazakhstan, Russia, and the Ukraine) are the worst performers in export diversification along the industrial dimension, while Belarus is the worst performer in geographical diversification among all transition countries. The above-mentioned lack of export diversification is even more striking when compared with the degree of export diversification of some of the regional leaders, such as the Czech Republic, Poland, and Slovenia, which have a substantially higher degree of export diversification relative to the levels predicted by the model.

Across different industries, the light industry, the timber wood and paper industry, the electronic equipment industry, the motor vehicles and parts industry, and other manufactured goods industry are the least geographically diversified. The highest geographical diversification is observed in industries that comprise the major share of total exports like the energy resources industry in Russia or the agriculture and metallurgy industries in the Ukraine.

The lack of diversification can be explained by some important deviations of export flows of the CIS countries from potential levels that result in lower degrees of geographical and industrial diversification. The CIS

countries tend to overtrade with other CIS countries and export disproportionately more in resource-extraction sectors. This is especially true in the energy resource industry, which particularly fails to fit the worldwide pattern. At the same time, the CIS countries export consistently less than expected in the agriculture, forestry, and food industries, which might indicate additional external and internal trade barriers that are particular to these industries. Central Asia and the Caucasus region have surprisingly weak trade connections with China and India and have great potential to increase their eastward exports. At the same time, Belarus, Moldova, and the Ukraine, underperform in their trade with the EU, especially in agriculture and food products.

The lack of diversification makes the CIS countries vulnerable to unfavorable terms of trade shocks. For example, out of all the transition countries, the Ukraine was hit the hardest by the current global economic crisis due to its reliance on metal exports, a commodity that suffered a devastating collapse in trade volumes. As another example, Belarus lacks geographical diversification of trade that makes this country vulnerable to trade disputes with Russia.

There is an important policy implication. Since, at the earlier stages of development, countries tend to diversify their industrial capacity and exports in different industries and markets for a sustainable economic growth, the CIS countries should further diversify their exports. Therefore, economic policy should be directed towards finding and promoting industries that have high

export potential, and it should refrain from supporting industries that extract and export raw materials.

APPENDIX: Model of Bilateral Export

Consider the Dixit-Stiglitz (1977) model of monopolistic competition with consumer preferences identical and homothetic across countries described, for example, by Feenstra (2003).¹⁴ Each country $i=1 \dots C$ has N_k^i firms that produce differentiated products in industries $k = 1, \dots, K$. Let c_{kl}^{ij} denote total consumption in country j of a good l that is produced by sector k in country i .

1. Consumers

A representative consumer located in country j has the utility function of the following form:

$$U^j = \sum_{k=1}^K \left(\int_{l \in B_k^j} c_{kl}^j \frac{\sigma-1}{\sigma} dl \right)^{\theta_k}, \quad (\text{A1})$$

where $\sigma > 1$ is the elasticity of substitution across different products. θ_k is the expenditure share of industry k in total consumption. B_k^j is the set of industry k goods that are available for consumption in country j .

The optimal consumption derived from the optimization problem is:

¹⁴ Chamberlin (1933) first introduced the main components of the monopolistic competition model. Dixit and Stiglitz (1977) and Krugman (1979) brought in the love of variety into the model.

$$c_{kl}^j = \frac{\theta_k Y^j}{P_k^j} \left(\frac{p_{kl}^j}{P_k^j} \right)^{-\sigma}, \quad (\text{A2})$$

where Y^j is the gross domestic product of country j that is equal to the total expenditures of country j .

$$P_k^j = \left(\int_{l \in B_k^l} (p_{kl}^j)^{1-\sigma} dl \right)^{\frac{1}{1-\sigma}} \quad (\text{A3})$$

is the price index of industry k .

2. Producers

A country i firm produces one unit of output with $w^i a$ units of labor.¹⁵ w^i is country specific, reflecting the differences in institutions, technology, and factor prices. a is a firm-specific parameter with the cumulative distribution function $G_k(a)$ over support $[a_{k \min}, a_{k \max}]$. Each firm is a monopolist over the production of a distinct good but is small relative to the size of the market. A standard formula for monopolistic pricing implies that the firm charges the mill price as a constant markup over the marginal cost:

¹⁵ We consider a partial equilibrium model with fixed capital during the period being investigated. Labor is the only input that is perfectly mobile across industries but immobile across countries.

$$p^i = \frac{\sigma}{\sigma-1} w^i a \quad (\text{A4})$$

There are variable and fixed costs of delivering products to consumer markets that vary across industries. T_k^{ij} is a melting iceberg transportation cost with $T_k^{ij} > 1, T_k^{ii} = 1$. F_k^{ij} is a fixed cost of exporting that is country-pair and industry specific with $F_k^{ij} > 0, F_k^{ii} = 0$. If the firm chooses to export its product to country j , consumers in country j pay $p_k^{ij} = \frac{T_k^{ij} \sigma w^i a}{\sigma-1}$. It follows that the profit of the firm exporting to country j is:

$$\pi_k^{ij}(a) = \frac{\theta_k}{\sigma} \left[\frac{\sigma T_k^{ij} w^i a}{(\sigma-1) P_k^j} \right]^{1-\sigma} Y^j - F_k^{ij}. \quad (\text{A5})$$

The firm exports only if it receives positive operating profits, which is more likely if the productivity of the firm ($\frac{1}{a}$) is high, the input price (w^i) is low, and the fixed costs of exporting (F_k^{ij}) are low. The least productive firm that exports to country j has the productivity level $1/a_k^{ij}$ determined as:

$$\pi_k^{ij}(a_k^{ij}) = 0 \Leftrightarrow \frac{\theta_k}{\sigma} \left[\frac{\sigma T_k^{ij} w^i a_k^{ij}}{(\sigma-1) P_k^j} \right]^{1-\sigma} Y^j = F_k^{ij}. \quad (\text{A6})$$

3. Industry Level Aggregation

Out of N_k^i firms that operate in country i in industry k , only $N_k^i G_k(a_k^{ij})$ firms export to country j . The aggregate export in industry k from exporter i to country j is:

$$X_k^{ij} = N_k^i G_k(a_k^{ij}) E(p_k^{ij}(a) c_k^{ij}(a) | \pi_k^{ij}(a) > 0) = N_k^i \theta_k Y^j V_k^{ij} \left(\frac{\sigma T_k^{ij} w^i}{(\sigma - 1) P_k^j} \right)^{1-\sigma}$$

$$\text{if } a_k^{ij} > a_{\min} \quad \text{and} \quad X_k^{ij} = 0 \quad \text{otherwise,}$$

$$\text{where } V_k^{ij} = \int_{a_{\min}}^{a_k^{ij}} a^{1-\sigma} dG_k(a).$$

The equation can be further simplified by using the equilibrium constraint on the output of sector k produced by country i that leads to the following export equation:

$$X_k^{ij} = s_k^i Y^i Y^j V_k^{ij} \frac{\left(\frac{T_k^{ij}}{P_k^j} \right)^{1-\sigma}}{\sum_{j=1}^C \left(\frac{T_k^{ij}}{P_k^j} \right)^{1-\sigma} Y^j V_k^{ij}} \quad (\text{A7})$$

References

- Acemoglu, Daron, Johnson, Simon, and Robinson, James A., 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review* 91 (December), 1369-1401.
- Athukurola, P.C., 2000, Manufacturing exports and terms of trade of developing countries: Evidence from Sri Lanka. *Journal of Development Studies* 36, 89-104
- Anderson, J. E., van Wincoop, E., 2003. Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* 93(1), 170-92.
- Babetskaia-Kukharchuk, O., Maurel, M., 2004. Russia's Access to the WTO: What Potential for Trade Increase? *Journal of Comparative Economics* 32(4), 680-699.
- Baldwin, Richard E. 1994. Towards an Integrated Europe. Centre for Economic Policy Research, London.
- Baldwin, R., Taglioni, D., 2006. Gravity for dummies and dummies for gravity equations. NBER Working Paper #12516.
- Bleaney, M., Greenaway, D., 2001. The impact of terms of trade and real exchange rate volatility on investment and growth in Sub-Saharan Africa. *Journal of Development Economics* 65, 491-500
- Chamberlin, Edward Hastings, 1933. *Theory of Monopolistic Competition*, Eighth Edition.
- Dixit, A., Stiglitz, J., 1977. Monopolistic Competition and Optimum Product Diversity. *American Economic Review* 67, 297-308.

- Egger, P., 2000. A note on the proper econometric specification of the gravity equation. *Economic Letters* 66, 25-31.
- Egger, P., 2002. An Econometric View of the Estimation of Gravity Models and the Calculation of Trade Potentials. *World Economy* 25(2), 297-312.
- Feenstra, R., 2003. *Advanced International Trade: Theory and Evidence*. Princeton University Press, Princeton and Oxford.
- Feenstra, R., Kee, Hiau Looi, 2008. Export variety and country productivity: Estimating the monopolistic competition model with endogenous productivity. *Journal of International Economics*, Volume 74, Issue 2 (March), 500-518
- Felbermayr, G, Kohler, W., 2006. Exploring the Intensive and Extensive Margins of World Trade, *Review of World Economics* 142(4), 642-674.
- Gros, D., Gonciarz, A., 1996. A note on the trade potential of Central and Eastern Europe. *European Journal of Political Economy* 12(4), 709-721.
- Hamilton, Carl B., and L. Alan Winters. 1992. Opening Up International Trade with Eastern Europe. *Economic Policy: A European Forum* (April), 77.
- Hausman, J. A., Taylor W. E., 1981. Panel Data and Unobservable Individual Effect. *Econometrica* 49, 1377-1398.
- Hausmann, Ricardo, Rodrik, Dani, 2003. Economic Development as Self-Discovery. *Journal of Development Economics* 72 (December), 603-633.

- Hausmann, Ricardo, Hwang, Jason, Rodrik, Dani, 2007. What You Export Matters. *Journal of Economic Growth* 12, No. 1, 1-25.
- Havrylyshyn, O., Al-Atrash, H., 1998. Opening Up and Geographic Diversification of Trade in Transition Economies. IMF Working Paper No. 98/22, Available at SSRN: <http://ssrn.com/abstract=882252>.
- Helpman, E., Melitz, M., Rubinstein, Y., 2008. Estimating Trade Flows: Trading Partners and Trading Volumes. *Quarterly Journal of Economics* 123(2), 441-487.
- Herzer, D., Danzinger, F. N-L., 2006. What does export diversification do for growth? An econometric analysis. *Applied economics* 38(15). 1825-1838.
- Hummels, David, 1999. Toward a Geography of Trade Costs. Available at SSRN: <http://ssrn.com/abstract=160533>.
- Hummels, David, and Klenow, Peter J., 2005. The Variety and Quality of a Nation's Exports. *American Economic Review* 95(3): 704-723.
- Imbs, Jean, Wacziarg, Romain, 2003. Stages of Diversification. *American Economic Review* 93 (March), 63-86.
- Kaufmann, D., Kraay, A., Mastruzzi, M., 2007. Governance Matters VI: Governance Indicators for 1996-2006. World Bank Policy Research Working Paper No. 4280, Available at SSRN: <http://www.ssrn.com/abstract=999979>.
- Koren, Miklos, Tenreyro, Silvana, 2007. Volatility and Development. *Quarterly Journal of Economics* 122 (February), 243-287.

- Krugman P., 1979. Increasing Returns, Monopolistic Competition, and International Trade. *Journal of International Economics* 9, 469-79.
- Limão, N., Venables, A., 2001. Infrastructure, Geographical Disadvantage, Transport Costs, and Trade. *The World Bank Economic Review* 15(3), 451.
- Lederman, Daniel, Maloney, William F., 2003. Trade Structure and Growth. World Bank Policy Research Working Paper No. 3025 (April). Available at SSRN: <http://www.ssrn.com/abstract=402460>.
- Martin, W., Pham, C., 2008. Estimating the Gravity Equation when Zero Trade Flows are Frequent. unpublished manuscript.
- McPherson, M., Trumbull, W., 2008. Rescuing Observed Fixed Effects: using the Hausman-Taylor Method for Out-of-Sample Trade Projections. *The International Trade Journal* 22(3), 315-340.
- Mehlum, Halvor, Moene, Karl, Torvik, Ragnar, 2006. Institutions and the Resource Curse. *Economic Journal* 116 (January), 1-20.
- Melitz, M., 2003. The impact of trade on intra-industry reallocation and aggregate industry productivity. *Econometrica* 71(6), 1695-1725.
- Pelzman, Joseph, 1977. Trade Creation and Trade Diversion in the Council of Mutual Economic Assistance: 1954-70. *American Economic Review* 67 (September), 713-722.
- Robinson, James A., Torvik, Ragnar, Verdier, Thierry, 2006. Political Foundations of the Resource Curse. *Journal of Development Economics* 79 (April), 447-468.

- Serlenga, L., Shin, Y., 2007. Gravity Models of Intra-EU Trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time specific factors. *Journal of Applied Econometrics* 22(2), 361-381.
- Silva, J., Tenreyro, S., 2006. The Log of Gravity. *The Review of Economics and Statistics* 88(4), 641-658.
- Wang, Z., Winters, L., 1991. The Trading Potential of Eastern Europe. CEPR Discussion Paper No. 610. London, Centre for Economic Policy Research. Available at <http://www.cepr.org/pubs/dps/DP610.asp>.
- World Bank, 2005. From Disintegration to Reintegration: Eastern Europe and the Former Soviet Union in International Trade. Washington, D.C. Available at <http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/ECAEXT/0,,contentMDK:20723133~pagePK:146736~piPK:146830~theSitePK:258599,00.html>.

Graphs and Figures

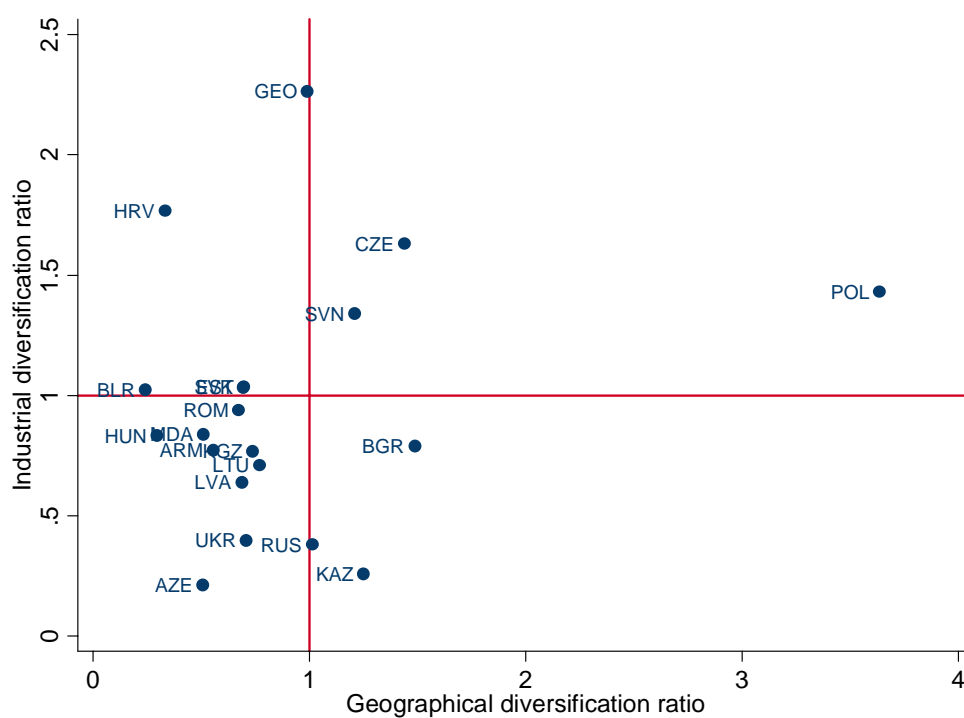


Figure 1. Export diversification of CIS and Eastern European countries in transition.

Table 1 Concordances between GTAP sectors and industries of the model

GTAP sector	Description	Sector of the model	Description
1	Paddy rice	1	Agriculture and forestry
2	Wheat	1	Agriculture and forestry
3	Cereal grains nec	1	Agriculture and forestry
4	Vegetables, fruit, nuts	1	Agriculture and forestry
5	Oil seeds	1	Agriculture and forestry
6	Sugar cane, sugar beet	1	Agriculture and forestry
7	Plant-based fibers	1	Agriculture and forestry
8	Crops nec	1	Agriculture and forestry
9	Bovine cattle, sheep and goats, horses	1	Agriculture and forestry
10	Animal products nec	1	Agriculture and forestry
11	Raw milk	1	Agriculture and forestry
12	Wool, silk-worm cocoons	1	Agriculture and forestry
13	Forestry	1	Agriculture and forestry
14	Fishing	1	Agriculture and forestry
15	Coal	2	Energy resources
16	Oil	2	Energy resources
17	Gas	2	Energy resources
18	Minerals nec	2	Energy resources
19	Bovine meat products	3	Food industry
20	Meat products nec	3	Food industry
21	Vegetable oils and fats	3	Food industry
22	Dairy products	3	Food industry
23	Processed rice	3	Food industry
24	Sugar	3	Food industry
25	Food products nec	3	Food industry
26	Beverages and tobacco products	3	Food industry
27	Textiles	4	Light industry
28	Wearing apparel	4	Light industry
29	Leather products	4	Light industry
30	Wood products	5	Timber, wood, pulp and paper
31	Paper products, publishing	5	Timber, wood, pulp and paper
32	Petroleum, coal products	6	Chemicals and petrochemicals
33	Chemical, rubber, plastic products	6	Chemicals and petrochemicals
34	Mineral products nec	6	Chemicals and petrochemicals
35	Ferrous metals	7	Metallurgy
36	Metals nec	7	Metallurgy
37	Metal products	7	Metallurgy
38	Motor vehicles and parts	8	Motor vehicles and parts
39	Transport equipment nec	8	Motor vehicles and parts
40	Electronic equipment	9	Electronic equipment
41	Machinery and equipment nec	10	Manufactures nec
42	Manufactures nec	10	Manufactures nec

Table 2 Export summary statistics

Industry	Average export, thousands of current \$US	Number of observed positive exports	Share of number of positive exports to number of potential exports
Agriculture and forestry	24620	43045	0.46
Energy resources	116770	27227	0.29
Food industry	43563	48522	0.52
Light industry	57722	48113	0.51
Timber, wood, and paper	35380	48858	0.52
Chemicals and petrochemicals	137636	54514	0.58
Metallurgy	73916	47441	0.51
Motor vehicles and parts	135227	41517	0.44
Electronic equipment	147949	41671	0.45
Manufactures nec	144730	54998	0.59
Overall	91264	455906	0.49

Note: The estimation sample has 128 exporters and 147 destinations in 10 industries in 2002-2006. Number of potential positive export links is calculated under assumption that within an industry all source countries trade with all destination countries.

Table 3 Definition of variables and data sources

Variables	Description	Sources
Dependent variables		
Export	Export from i to j in sector k, in thousands of current \$US. COMTRADE exports data aggregated to GTAP sectors are calculated based on HS1996 classification in 2002-2006 – CIS excluded – and further aggregated to industries of the model.	United Nations Commodity Trade Statistics Database
Independent variables		
Sector export share	(Total export of country i in industry k - export from i to j in industry k)/Total export of country i	Author's calculations
GDP	Gross domestic product, in current \$US.	World development indicators
Population	Population	World development indicators
Interior distance	Internal distance of country, (an often used measure of average distance between producers and consumers in a country, see Head and Mayer, 2002, "Illusory Border Effects", CEPII Working Paper No. 2002-01, for more on this topic).	CEPII
Dist	$d_{ii} = 0.67\sqrt{area/\pi}$ distance between the biggest cities of countries i and j. d_{kl} is the distance between cities k and l. (Head and Mayer, 2002) $d_{ij} = \sum_{k \in i} (pop_k / pop_i) \sum_{l \in j} (pop_l / pop_j) d_{kl}$	CEPII
Landlocked	Dummy variable set equal to 1 for landlocked countries.	CEPII
Contig	Dummy variables indicating whether the two countries are contiguous.	CEPII
Remoteness	log of average GDP of all other countries weighted by inverse of the distance to those countries $\ln MR_i^j = \ln \left(\sum_{j=1}^C \frac{GDP_j^j}{dist_{ij}^j} \right)$	Author's calculations
Colony	Dummy variable set equal to 1 if one of the countries used to be a colony of the other country.	CEPII
Same continent	Dummy variable set equal to 1 if countries i and j located on the same continent. Mapping of countries to continents was taken from CEPII	Author's calculations
Selection variables		
Common language	Dummy variable indicating whether countries share a common language.	CEPII
Reg. quality	Regulatory quality index measures the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development (Kaufmann, Kraay and Mastruzzi, 2007)	Governance matters, 2007

Table 4 Selection equation

	Agriculture and forestry	Energy resources	Food industry	Light industry	Timber, wood and paper	Chemicals and petrochemicals	Metalurgy	Motor vehicles and parts	Electronic equipment	Manufactures nec
Common lang. Yes=1	0.17*** (0.011)	0.093*** (0.012)	0.15*** (0.011)	0.13*** (0.011)	0.17*** (0.011)	0.11*** (0.010)	0.13*** (0.011)	0.14*** (0.012)	0.16*** (0.012)	0.15*** (0.009)
Reg. quality i	0.17*** (0.007)	0.15*** (0.006)	0.27*** (0.008)	0.19*** (0.007)	0.20*** (0.008)	0.17*** (0.007)	0.17*** (0.008)	0.21*** (0.008)	0.29*** (0.009)	0.20*** (0.007)
Log of distance	-0.13*** (0.007)	-0.22*** (0.006)	-0.16*** (0.007)	-0.17*** (0.007)	-0.20*** (0.007)	-0.21*** (0.007)	-0.22*** (0.007)	-0.23*** (0.007)	-0.17*** (0.007)	-0.19*** (0.007)
Log of GDP per capita i	0.019*** (0.004)	0.043*** (0.003)	0.059*** (0.004)	0.082*** (0.004)	0.10*** (0.004)	0.13*** (0.004)	0.15*** (0.004)	0.14*** (0.004)	0.12*** (0.005)	0.12*** (0.004)
Contig. Yes=1	0.23*** (0.035)	0.18*** (0.039)	0.17*** (0.036)	0.15*** (0.034)	0.11** (0.039)	0.014 (0.039)	0.14*** (0.040)	0.17*** (0.038)	0.20*** (0.034)	0.038 (0.036)
Same cont. Yes=1	0.12*** (0.011)	0.0078 (0.009)	0.12*** (0.011)	0.12*** (0.010)	0.11*** (0.011)	0.078*** (0.010)	0.076*** (0.011)	0.078*** (0.011)	0.13*** (0.012)	0.064*** (0.010)
Colony Yes=1	0.26*** (0.040)	0.26*** (0.041)	0.16*** (0.040)	0.19*** (0.038)	0.13** (0.041)	0.089* (0.043)	0.16*** (0.046)	0.22*** (0.045)	0.18*** (0.044)	0.073 (0.042)
Ln Int. dist i	0.19*** (0.004)	0.19*** (0.004)	0.20*** (0.004)	0.17*** (0.004)	0.19*** (0.004)	0.18*** (0.004)	0.21*** (0.004)	0.22*** (0.004)	0.21*** (0.004)	0.18*** (0.003)
Landlocked i. Yes=1	-0.12*** (0.009)	-0.087*** (0.006)	-0.18*** (0.010)	-0.14*** (0.009)	-0.14*** (0.009)	-0.12*** (0.009)	-0.070*** (0.009)	-0.076*** (0.009)	-0.061*** (0.009)	-0.030*** (0.008)
Chi-squared	9318.0	9661.8	9241.9	9224.0	9745.9	9092.9	10226.0	10268.2	10487.4	9899.9
Pseudo R-squared	0.31	0.41	0.34	0.31	0.35	0.36	0.38	0.38	0.41	0.37
Observations	88798	88798	88798	88798	88798	88798	88798	88798	88798	88798

* p<0.05, ** p<0.01, *** p<0.001

Note: (1) The dependent variable is a dummy variable which takes value of one if there is a positive trade flow from country i to country j and zero otherwise. The overall sample has 128 source countries and 147 destination countries in 2002-2006. A constant term, destination country time specific fixed effects are included but not reported. (2) Marginal effects of probit regression for each industry are presented. For dummy variables, marginal effects are calculated for discrete change from 0 to 1. Country-pair cluster-robust standard errors are reported in parentheses.

Table 5 Gravity model of export estimated by Hausman-Taylor method

	Agriculture and forestry	Energy resources	Food industry	Light industry	Timber, wood and paper	Chemicals and petrochemicals	Metalurgy	Motor vehicles and parts	Electronic equipment	Manufactures nec
Log of sector exp. share	0.73** (0.019)	0.41** (0.017)	0.75** (0.018)	0.65** (0.015)	0.70** (0.019)	0.41** (0.021)	0.47** (0.020)	0.54** (0.017)	0.55** (0.015)	0.60** (0.018)
Log of GDP i	0.99** (0.030)	0.84** (0.043)	1.07** (0.028)	1.33** (0.025)	1.25** (0.030)	1.31** (0.026)	1.40** (0.029)	1.44** (0.030)	1.20** (0.030)	1.26** (0.023)
Log of GDP j	0.42** (0.059)	0.27** (0.092)	0.37** (0.048)	0.39** (0.049)	0.41** (0.051)	0.44** (0.042)	0.70** (0.057)	0.59** (0.063)	0.52** (0.057)	0.51** (0.043)
Log of distance	-0.86** (0.049)	-1.46** (0.092)	-1.00** (0.047)	-1.05** (0.051)	-1.47** (0.053)	-1.64** (0.051)	-1.54** (0.054)	-1.03** (0.055)	-0.81** (0.059)	-1.07** (0.045)
Contig. Yes=1	2.00** (0.163)	0.73** (0.233)	1.88** (0.158)	2.04** (0.170)	1.44** (0.172)	1.62** (0.166)	1.57** (0.170)	1.50** (0.168)	1.59** (0.186)	1.56** (0.152)
Same cont. Yes=1	0.70** (0.079)	-0.000069 (0.125)	0.62** (0.076)	1.21** (0.080)	0.72** (0.081)	0.84** (0.077)	0.72** (0.082)	0.85** (0.084)	1.00** (0.094)	0.73** (0.070)
Colony Yes=1	1.13** (0.179)	0.77** (0.244)	1.13** (0.174)	1.21** (0.187)	1.24** (0.190)	0.87** (0.184)	1.12** (0.186)	0.85** (0.182)	0.98** (0.202)	0.98** (0.169)
Ln Int. dist i	0.18** (0.045)	0.24** (0.078)	0.057 (0.037)	-0.093** (0.035)	-0.15** (0.039)	-0.12** (0.033)	-0.24** (0.039)	-0.41** (0.044)	-0.27** (0.041)	-0.20** (0.031)
Landlocked i, Yes=1	-0.34** (0.087)	-0.90** (0.152)	-0.42** (0.089)	0.0024 (0.093)	-0.093 (0.097)	-0.56** (0.092)	0.13 (0.101)	0.51** (0.101)	-0.23* (0.103)	0.13 (0.075)
Remoteness i	-0.96** (0.273)	-1.52** (0.420)	-0.18 (0.251)	-0.39 (0.227)	-0.55* (0.272)	1.77** (0.236)	-0.032 (0.284)	-1.19** (0.315)	1.83** (0.301)	1.16** (0.223)
η	2.15** (0.243)	1.33** (0.277)	0.31 (0.197)	1.77** (0.237)	1.18** (0.202)	1.00** (0.210)	1.07** (0.218)	1.45** (0.217)	1.14** (0.187)	0.74** (0.180)
ψ	6.47** (0.896)	4.32** (1.208)	1.99** (0.691)	3.69** (0.880)	2.92** (0.708)	1.93** (0.682)	2.55** (0.762)	3.47** (0.834)	3.23** (0.721)	1.65** (0.603)
ψ^2	-2.15** (0.374)	-1.18* (0.533)	-0.58* (0.279)	-1.09** (0.367)	-0.74** (0.285)	-0.62* (0.267)	-0.72* (0.305)	-0.90** (0.344)	-0.88** (0.294)	-0.25 (0.238)
ψ^3	0.22** (0.050)	0.12 (0.074)	0.049 (0.036)	0.074 (0.049)	0.051 (0.037)	0.050 (0.034)	0.053 (0.039)	0.064 (0.045)	0.068 (0.038)	-0.0078 (0.030)
$\chi^2(4)$: $b_0=0, b_1=0, b_2=0, b_3=0$	129.11	112.32	64.22	136.07	102.63	65.86	59.67	118.33	95.77	119.51
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
χ^2	12323.5	5479.9	15111.0	16493.6	16631.9	21858.1	17777.5	16959.7	17236.8	28804.0
σ_u^2	2.37	3.13	2.38	2.56	2.64	2.58	2.56	2.48	2.79	2.39
σ_c^2	1.14	1.38	1.04	1.06	1.14	1.03	1.24	1.29	1.22	1.04
ρ	0.81	0.84	0.84	0.85	0.84	0.86	0.81	0.79	0.84	0.84
Observations	37532	23876	41763	41331	41814	45871	40025	35292	35744	46287

* p<0.05, ** p<0.01

Note: The dependent variable is log of export from country i to country j. The overall sample has 120 source countries and 147 destination countries in 2002-2006. 8 CIS countries are not included into estimation. Standard errors are reported in parentheses. Destination country fixed effects, time dummies, and constant term are included but not reported. Results of Hausman-Taylor regression by industries are presented. The endogenous variables that can be correlated with the random effect error term u^i is remoteness of i.

Table 6 Geographical patterns of CIS exports

Reporter	Partner					
	China	CIS	EU	India	RoW	USA
Armenia	0.011	0.177	0.444	0.000	0.287	0.080
	0.035	0.141	0.224	0.021	0.484	0.096
Azerbaijan	0.009	0.151	0.595	0.001	0.230	0.015
	0.036	0.399	0.162	0.022	0.318	0.063
Belarus	0.023	0.481	0.422	0.007	0.050	0.016
	0.005	0.146	0.730	0.002	0.066	0.051
Georgia	0.007	0.466	0.229	0.008	0.249	0.042
	0.018	0.609	0.173	0.007	0.143	0.051
Kazakhstan	0.098	0.172	0.373	0.001	0.341	0.014
	0.376	0.295	0.105	0.041	0.124	0.060
Kyrgyzstan	0.054	0.394	0.048	0.001	0.487	0.015
	0.511	0.142	0.130	0.047	0.097	0.073
Moldova	0.001	0.498	0.419	0.001	0.045	0.037
	0.002	0.129	0.750	0.001	0.065	0.052
Russia	0.063	0.111	0.576	0.010	0.207	0.033
	0.221	0.110	0.357	0.025	0.195	0.091
Ukraine	0.024	0.294	0.335	0.016	0.297	0.034
	0.010	0.180	0.654	0.005	0.098	0.054

Note: For each country-pair the number on top is the actual share of export from country reporter to country partner divided by overall export to partners and the number at the bottom is the share predicted by the gravity equation.

Table 7 Industrial structure of exports in CIS

Reporter	Industry									
	Agriculture and forestry	Energy resources	Food industry	Light industry	Timber, wood and paper	Chemicals and petrochemicals	Metallurgy	Motor vehicles and parts	Electronic equipment	Manufactures nec
Armenia	0.01	0.10	0.11	0.05	0.01	0.04	0.29	0.00	0.00	0.39
	0.03	0.02	0.17	0.07	0.01	0.08	0.25	0.01	0.01	0.36
Azerbaijan	0.04	0.64	0.03	0.00	0.00	0.21	0.04	0.03	0.00	0.01
	0.17	0.04	0.12	0.03	0.01	0.30	0.17	0.09	0.00	0.06
Belarus	0.01	0.03	0.08	0.07	0.05	0.45	0.08	0.07	0.01	0.16
	0.03	0.01	0.10	0.11	0.05	0.45	0.07	0.03	0.01	0.15
Georgia	0.09	0.09	0.22	0.01	0.02	0.10	0.27	0.14	0.00	0.05
	0.24	0.02	0.44	0.02	0.02	0.09	0.08	0.03	0.00	0.04
Kazakhstan	0.03	0.66	0.01	0.01	0.00	0.06	0.21	0.01	0.00	0.01
	0.15	0.13	0.09	0.04	0.00	0.15	0.34	0.03	0.01	0.07
Kyrgyzstan	0.14	0.01	0.06	0.07	0.01	0.21	0.41	0.03	0.00	0.06
	0.35	0.02	0.09	0.08	0.01	0.13	0.23	0.02	0.00	0.08
Moldova	0.18	0.00	0.39	0.22	0.02	0.06	0.04	0.01	0.00	0.06
	0.22	0.00	0.35	0.15	0.02	0.13	0.03	0.01	0.01	0.07
Russia	0.03	0.44	0.01	0.01	0.03	0.24	0.18	0.03	0.00	0.03
	0.07	0.15	0.05	0.03	0.06	0.30	0.19	0.04	0.01	0.10
Ukraine	0.05	0.05	0.08	0.04	0.03	0.17	0.42	0.06	0.01	0.09
	0.11	0.02	0.16	0.09	0.04	0.24	0.18	0.04	0.01	0.12

Note: For each country-industry pair the number on top is the actual share of export and the number at the bottom is the predicted share.

Table 8 Geographical concentration of CIS exports

Reporter	Industry									
	Agriculture and forestry	Energy resources	Food industry	Light industry	Timber, wood, and paper	Chemicals and petrochemicals	Metallurgy	Motor vehicles and parts	Electronic equipment	Manufactures nec
<i>Armenia</i>	0.83	0.42	0.18	0.62	0.63	0.76	0.77	1.10	0.42	0.34
<i>Azerbaijan</i>	0.68	0.28	1.06	0.23	0.69	1.38	0.67	0.23	0.46	1.34
<i>Belarus</i>	0.44	0.29	0.10	0.31	0.26	0.48	0.34	0.22	0.18	0.19
<i>Georgia</i>	2.89	0.32	1.97	0.81	0.56	1.21	0.62	0.77	1.46	1.27
<i>Kazakhstan</i>	1.41	3.25	4.09	0.35	0.82	3.61	2.20	1.39	1.20	0.52
<i>Kyrgyzstan</i>	0.97	1.43	0.40	0.66	0.55	1.15	0.95	0.91	0.35	0.86
<i>Moldova</i>	0.92	0.75	0.19	0.63	0.36	0.39	0.69	0.43	0.81	0.45
<i>Russia</i>	0.45	2.31	0.54	1.04	1.52	1.34	2.22	0.70	0.66	0.77
<i>Ukraine</i>	2.41	0.89	1.03	0.72	0.46	1.49	1.54	0.38	0.46	0.42
<i>CIS average</i>	1.22	1.10	1.06	0.59	0.65	1.31	1.11	0.68	0.67	0.68

Note: For each country and industry represented in the table, the HHI of geographical concentration is computed for the actual and predicted exports. The index is normalized and takes values between 0 and 1 with higher numbers representing higher concentration of exports for certain destinations. The table reports the ratio of the predicted HHI to the actual HHI. When the ratio takes value above 1 it means that actual exports are more diversified than the predicted ones while the ratio below 1 indicates that the predicted diversification is higher than the actual one.

Table 9 Industrial concentration of CIS exports

Reporter	Year					
	2002	2003	2004	2005	2006	Average 2002-2006
<i>Armenia</i>	0.73	0.70	0.79	0.84	0.81	0.77
<i>Azerbaijan</i>	0.07	0.13	0.24	0.33	0.28	0.21
<i>Belarus</i>	1.14	1.10	1.04	0.94	0.90	1.02
<i>Georgia</i>	1.86	2.21	2.21	2.39	2.65	2.26
<i>Kazakhstan</i>	0.33	0.27	0.23	0.22	0.23	0.26
<i>Kyrgyzstan</i>	1.07	0.66	0.67	0.67	0.77	0.77
<i>Moldova</i>	0.78	0.77	0.81	0.85	0.98	0.84
<i>Russia</i>	0.38	0.34	0.37	0.40	0.41	0.38
<i>Ukraine</i>	0.38	0.46	0.41	0.39	0.35	0.40
<i>CIS average</i>	0.75	0.74	0.75	0.78	0.82	0.77

Note: For each country and year represented in the table, the HHI of industrial concentration is computed for the actual and predicted exports. The index is normalized and takes values between 0 and 1 with higher numbers representing higher concentration of exports in certain industries. The table reports the ratio of the predicted HHI to the actual HHI. When the ratio takes value above 1 it means that predicted exports are less diversified than the actual ones while the ratio below 1 indicates that the predicted diversification is higher than the actual one.